Origin and dispersal of *Homalictus* (Hymenoptera: Halictidae): Phylogenetics of a halictine bee sub-genus

Scott Groom

Biological Sciences, Flinders University

GPO BOX 2100, Adelaide, SA 5001, Australia

Phone: +61 8 8201 2540

Fax: +61 8 8201 3015

Email: scott.groom@flinders.edu.au

Address for correspondence:
Scott Groom

School of Biological Sciences

Flinders University

GPO Box 2100

Adelaide SA 5001

Abstract

Halictine bees are important in studies of insect sociality due to the substantial diversity in social behaviour among species. All Australian halictines studied to date are known to be communal, suggesting they arose through a single invasion. *Homalictus* is a subgenus of the halictines, widely distributed throughout Southeast Asia and Australia. The origin of halictines in Australia, and thus a minimum time for the origin of communal living, is uncertain. Here, previous studies are used as a basis to further analyse halictine relationships and subsequently estimate the time of origin and divergence of *Homalictus*. It is shown that the origin of this sub-genus is older than previously anticipated and communal behaviour has been stable over long periods of evolutionary time. Due to the estimated 33 million years before present (myBP) origin of *Homalictus* within the Oligocene (34-23 myBP), recent routes of halictine dispersal into Australia from Asia are difficult to explain. Future studies including representatives of neighbouring continents are needed to explore the possibility of dispersal from Africa.

Key words: Phylogenetics, Halictinae, halictine, *Lasioglossum*, *Homalictus*, Social Evolution, Communal, Australia

1. Introduction

Halictine bees (subfamily Halictinae) provide an ideal group for phylogenetic studies due to the substantial diversity in social behaviour among species (Danforth 2002), ranging from strictly solitary to highly eusocial with every intervening grade of sociality known. Halictine bees have been a major focus in studies of social evolution since the 1960s (Michener 1958; Batra 1966; Smith and Wenzel 1988; Wcislo et al. 1993; Danforth 1999; Danforth and Ji 2001; Danforth et al. 2003; Schwarz et al. 2007). Their unique range of social forms has allowed this group of bees to become a major taxon in studies of insect sociality (Knerer and Schwarz 1976; Danforth and Ji 2001; Danforth et al. 2003; Brady et al. 2006; Schwarz et al. 2007).

Previous hypotheses concerning the history of social evolution in halictine bees have not been based on explicit phylogenies, and thus are highly speculative (Danforth 2002). Recent studies have begun analysing the phylogenetic relationships of halictine bees based on both nuclear and mitochondrial DNA sequence data. This has allowed insights into lower level (generic and subgeneric) relationships to be more confidently inferred. Understanding the phylogeny, antiquity, and early radiation of these bees is important because of their importance for understanding the timing of pollinator/angiosperm coevolution (Danforth et al. 2004). Recent molecular studies have allowed important insights into halictine evolution (Schwarz et al. 2007) that were not evident from earlier, morphology-based studies such as those of Michener (1965; 1979).

Establishing confident reconstructions of historical biogeographical movements of bees based upon morphology has so far been near impossible. Michener (1965; 1979) proposed hypotheses of biogeographic movement for many bee groups, but phylogenetic evidence was based on morphology alone. Identification of some halictine species may be impeded by an overlap in defining characteristics due to high morphological variability, especially colouration. Knerer and Schwarz (1976) described the difficulty involved in using living forms to confidently infer evolutionary pathways of insect societies due to such confounding variation.

Australian halictine species display a unique array of behavioural attributes that distinguish them from species of other parts of the world (Michener 1960; Knerer and Schwarz 1976). All Australian halictines studied to date exhibit either solitary or communal nesting behaviour (Michener 1960), a situation in which multiple females share a nest but do not cooperate in cell provisioning or show reproductive division of labour (Danforth and Ji 2001). Outside of Australia communal behaviour is rare in halictines and extremely rare in insects as a whole. Knerer and Schwarz (1976) termed this unique range of behaviour characteristics the 'Australian Enigma'. Despite their importance, research on halictine sociality has been hindered by poor understanding of phylogenetic relationships at the generic and species levels (Danforth 1999).

Danforth and Ji (2001) resolved the 'Australian Enigma' by showing that the Australian halictines form a monophyletic clade. This suggests that Australian halictines must be the result of a single colonisation event by a communal, ancestral halictine and all subsequent species have retained this sociality type. Communal

nesting would consequently reflect common ancestry rather than convergent evolution. Aside from sociality type, Australian *Lasioglossum* including the sub-genus *Homalictus,* are known to share flower associations and nest architecture (Knerer and Schwarz 1976). Danforth and Ji (2001) suggest an estimated date of Australian origin circa 30 myBP. However, due to a minor representation of Australian species (in particular *Homalictus*) the power of their study is limited. The phylogenetics of their study contained a polytomy that included the Australian species and *Homalictus* clade, and failed to include a formally estimated time of origin of *Homalictus*.

Recent phylogenetic studies, such as Danforth et al. (2004), used molecular data to establish broad relationships for higher level bee groups (family, tribe) with strong support from combinations of nuclear and mitochondrial genes. Danforth et al. (2004) proposed a Southern Hemisphere origin for halictine bees, with subsequent dispersal events into the Northern Hemisphere. They estimated dispersal of approximately 30 myBP for halictine bees into Australia. However, their study included few Australian species and no Asian halictines, which could comprise potential ancestral clades giving rise to the Australian clade, so the possibility of a Laurasian dispersal route could not be evaluated. Brady et al. (2006) found that the three origins of eusociality in halictines were both recent and simultaneous (~22 myBP), possibly correlated temporally with a period of global warming. They suggested the climatic influences of this period may have influenced evolution of this sociality type, in addition to diversification in halictine groups.

The suggested origin of eusociality in halictine bees (20-22 myBP) by fossil-calibrated molecular analyses is far more recent than the estimated minimum 65 myBP origins of sociality in corbiculate bees, vespid wasps, ants and termites (Brady et al. 2006). This raises the question of convergent factors that may have been present for origins of sociality to occur. Brady et al. (2006) suggested climatic factors, however, these cover a large spatial area and would not explain the impediment of sociality arising in communal lineages. In contrast to three known origins of eusociality, Danforth (2002) showed that there have been as many as 12 losses in the halictines, indicating that eusociality is hard to attain but easily lost. Geographic location may have influenced the competitive pressure associated with bee faunas in certain areas. As predominantly generalist foragers, halictines in Australia may not have been exposed to the selective pressures of Northern Hemisphere bee fauna due to the availability of vacant niches (Danforth et al. 2004).

Brady et al. (2006) showed much of the Halictinae diversification occurred within the Tertiary Epochs of the Oligocene (33.9-23.0 myBP) and Miocene (23.0-5.33 myBP) (Gradstein et al. 2004). Geological data from the Late-Oligocene to Early-Miocene indicates that there was considerable movement of tectonic plates within the Indo-Australian region, particularly Sulawesi (Hall 2002). The Early-Oligocene saw the movement of Borneo and the Philippines southward and by the Late-Oligocene there was a potential corridor for movement between the Australian and Laurasian plates (Hall 2002). This ‘corridor’, although consisting of several fragmented islands, may have facilitated ‘island hopping’. In terms of biogeographical movement, timing in the Oligocene would determine the size of obstruction that has been overcome for colonisation to occur. The Laurasian-Australian interchange is thought to have

occurred within the Miocene (de Jong 2004). It is therefore quite plausible for a single colonisation of Australia to have occurred from the Early-Miocene, as suggested by Danforth and Ji (2001), with subsequent back migrations during later periods. Higher dispersal rates between Australia and the islands of Southeast Asia would be expected during these later periods when islands were within close proximity.

Danforth et al. (2001) showed that Australian halictines share a common ancestor, which suggests the communal behaviour exhibited by all subsequent species has been retained through diversification. This highlights two main issues, (i) as the work of Danforth et al. (2001) and subsequent studies have only minimal representation of the Australian species in phylogenetic analyses estimating the timing of halictine origin, the origin of communal living in Australia is uncertain. (ii) The timing and source location of the colonising halictine into Australia has not been explicitly investigated. Previous studies have lacked representatives of Asia and Africa in biogeographical reconstructions, thus comparisons of dispersal cannot be made accurately. The idea of 'when' these origins occurred is linked through this ancestral halictine and can be investigated simultaneously as the ancestral halictine timing infers the origin of communal behaviour in Australia.

Homalictus is a *Lasioglossum* subgenus of the tribe Halictini. For the purposes of this paper, the genus *Lasioglossum* and its numerous subgenera will initially be referred to as *L. (Homalictus)* for the subgenus *Homalictus*, thereafter referred to generally as *Homalictus* or abbreviated for species to *L. (H.) sp*. Walker (1997) revised the subgenus to 44 species, including five species previously undescribed in the study of Walker (1986). *Homalictus* is represented throughout the Indo-Pacific region from Sri

Lanka through the Philippines to Samoa, and south to Australia. The region is highly fragmented, consisting of islands and areas of high elevation. The landscape provides substantial geographic barriers, which have the potential to drive selection. Danforth and Ji (2001) suggested that the presence of species of *Homalictus* outside of Australia has probably resulted from dispersal from Australia. Much of *Homalictus* diversity in Australia occurs in eastern coastal areas, but southern Western Australia also exhibits a high degree of diversity with the potential for increase due to previously limited sampling in the area. Therefore, poor knowledge of taxonomy and distributions across Australia may limit the known extent of *Homalictus* in Australia. Studies such as Hingston (1998) have expressed difficulty in differentiating *Homalictus* species. The similarity in morphology of (*L.) Homalictus niveifrons* (Cockerell, 1914) and *L. (H.) megastigmus* (Cockerell, 1926) resulted in the two being combined in their analyses. Previous studies suggest that *Homalictus* diverged between 5-10 myBP, however, this estimation is based on a chronogram where taxa were selected primarily to reveal earlier divergences in the Halictinae (Danforth et al. 2004; Brady et al. 2006). Consequently the support for the *Homalictus* node is limited. The two *Homalictus* species of Danforth et al. (2004) produced the only incongruence among the three genes used in their study. Due to the limited number of Australian species included there is substantial potential for error in estimation of Australian halictine divergence times.

Donovan (1983) showed that the distance dispersed by *Homalictus* when it colonised New Caledonia is considerably higher compared to other halictines and colletid bees. He suggests *Homalictus* species of New Caledonia were the outcome of a single colonisation by an Australian *Homalictus* following the rifting of New Caledonia

from eastern Australia in the Late-Cretaceous (Yan and Kroenke 1993; Barker et al. 2007). Michener (1979) also suggests that small ground-nesting bees such as *Homalictus* disperse across oceans more readily than the larger species.

This study uses phylogenetic methods to investigate the origin and dispersal of the halictine bee sub-genus, *Homalictus*. Further analyses using dating techniques will investigate the time of origin and implications of this on the origin of Australian halictines communal behaviour. Representation of Australian halictines in previous studies has been minimal (Danforth and Ji 2001; Danforth et al. 2004; Brady et al. 2006). This study will utilise available resources to further the work of others with emphasis on the represented Australian clades.

2. Methods

2.1. Taxa Collection

Overall, 73 putative taxa were collected (Appendix A) from across southern to south western and northern to north eastern Australia. In addition to this, collections were also made on the island of Borneo (Sarawak), Malaysia with limited success. Specimens were killed using ethyl acetate and transferred directly to 99% ethanol for molecular preservation. Ethanol was changed three times within the first week to minimise DNA degradation. Additional sequences were acquired through GenBank (Benson et al. 2006) (see Appendix B).

2.2. DNA Extraction, Amplification and Sequencing.

DNA extractions followed Gentra Puregene Cell Kit (Qiagen) standard protocols. Tissue samples of approximately 5 mg were taken from three legs of each specimen. Macerated tissue was transferred to 1.5 μl microfuge Eppendorf tubes containing 300 μl cell lysis solution and 1.5 μl of protenase K, and incubated overnight at 55°C.

Samples were then cooled to room temperature, vortexed for 10 seconds, placed on ice for 10 minutes, and centrifuged for 5 minutes. To precipitate proteins, 100 μl of ammonium acetate was then added and mixed. Samples were again placed on ice and centrifuged at 15 000 x g to form a protein pellet.

The supernatant was then transferred by pipette to a clean 1.5 μl Eppendorf microfuge tube containing 300 μl of 100 % isopropanol and 1.0 μl of Glycogen solution (20 mg/ml). Samples were then inverted gently 50 times and stored overnight at room temperature. Following centrifuging at 15 000 x g for 10 minutes, the supernatant was removed by pipette, leaving the DNA pellet. 300 μl of 70 % ethanol was used to wash the pellet. The samples were again centrifuged at 15 000 x g for 5 minutes, and the ethanol removed and left to dry. DNA hydration was achieved by the addition of 50 μl of TLE and storage overnight at 4°C.

Polymerase chain reaction (PCR) processes utilised Hot Master Taq (Promega) and were based on standard protocols of Danforth (1999). One mitochondrial gene and one nuclear gene region were amplified and sequenced bi-directionally (Schwarz et al. 2006). The mitochondrial region was from the protein coding gene Cytochrome Oxidase I (COI) and the nuclear exon region was from the F2 copy of Elongation Factor - 1α (EF-1α). The COI primer combination of Jerry/Pat (Danforth 1999) produced a consistently amplified 900-bp PCR product that was sequenced accurately by automated sequencing. Cycle conditions were as follows: 94°C, 1 min denaturation; 50°C, 1 min annealing; 72°C, 1 min 15 s extension for a total of 34 cycles (Danforth 1999). The EF1-α F2 copy specific combination of HaF2For1/F2-rev1 was used to produce an approximately 1100-bp fragment (Danforth et al. 1999). This primer combination is commonly used in halictine studies (Danforth and Ji 2001; Danforth et al. 2003; Danforth et al. 2004; Brady et al. 2006). Cycle conditions were as follows: 94°C, 1 min denaturation; 54°C, 1 min annealing; 72°C, 1 min 30s extension for a total of 35 cycles (Danforth et al. 1999).

PCR products were purified directly using magnetic bead manual clean up (AMPure) or Multiscreen PCR_{384} Filter Plate (Millipore). PCR purification was achieved through circa 50 µl of product sequenced in 20 µl reaction volumes for each original PCR primer using the Big Dye Ready Reaction kit Version 3.1 (Applied Biosystems). Reaction products were then purified by isopropanol precipitation or Millipore Filter plate, which both provided consistently high quality sequences. Sequencing PCR products were sent to IMVS, Adelaide or Macrogen DNA sequencing services, Korea, for automated sequencing. In total 73 specimens were sequenced successfully for the two genes.

Forward and reverse sequence chromatograms were compared using SeqEd 1.03 (Applied Biosystems) to reveal inconsistencies. Consensus sequences were then exported and aligned with SeAl Sequence Alignment Editor (Rambaut and Charleston 2002) software. The European honeybee, *Apis mellifera* was included as a reference to determine the reading frame of each sequence, but was subsequently removed for phylogenetic analyses. The intron region of EF-1α was excluded from analyses due to large fragment of bases which were impossible to align (Schwarz et al. 2006). In total 123 sequences including 50 from studies of Danforth et al. (1999; 2004) (see Appendix B for accession numbers) were used in analyses.

2.3. Phylogenetic Analyses

Phylogenetic analyses were designed to address two main issues, (i) the timing of halictine origin, and thus the origin of communal living in Australia. Although this has been addressed by Danforth et al. (2001) and subsequent studies, representation of

Australian halictine species was minimal in phylogenetic analyses; thus interpretations of analyses cannot be made with confidence. (ii) The timing and source location of the colonising halictine has not been explicitly investigated.

2.3.1 Location of Origin and Subsequent Dispersals

In order to recover the phylogeny of *Homalictus* both maximum parsimony (MP) and Bayesian inference (BI) were used. For bees, BI is regarded as a more reliable approach due to the inability of MP in accounting for heterogeneity in transition matrices among gene partitions and signal degradation commonly associated with the third mitochondrial codon position. The extreme AT bias exhibited at 3rd mtDNA positions leads to substantial problems in phylogenetics due to the high rate of homoplasic change (Schwarz et al. 2004). This was confirmed for the *Homalictus* dataset as seen in Table 1 (Results). Single models applied to a combined data set fail to account for difference in both base composition and evolutionary rates between partitions (Schwarz et al. 2006). BI is a more suitable method due to its ability to fit an appropriate model to the data, and so, adequately account for the variability at codon positions and data partitioning. Dispersal was therefore largely inferred from the results of Bayesian analyses (MrBayes version 3.1.2 (Huelsenbeck and Ronquist 2001)).

A dataset containing both, EF-1α and COI, gene regions for all *Homalictus* specimens was constructed. *Agapostemon tyleri* (**AF140320, AF102835**) of the Augochlorini tribe was included as an outgroup species. Also included to reduce long branch length attraction were *L. (Dialictus) zephyrum* (**AF435379, AF103973**)*, L. (Parasphecodes)*

hybodinum (**AF264857, AF104660**), *L. (Chilalictus) conspicuum* (**AF264789, AF103952**), *L. (Chilalictus) lanarium* (**AF264793, AF103956**), *L. (Chilalictus) mediopolitum* (**AF264794, AF103957**), and *L. (Chilalictus) florale* (**AF264792, AF103955**), obtained through GenBank (Benson et al. 2006).

MP analyses involved ten random stepwise sequence additions holding two trees at each step and with tree bisection and reconnection (TBR) for searching tree space. 500 bootstrap pseudo-replicates were used to assess support for nodes. MP analyses have generally been unreliable in bee phylogenetic studies because of the age of many nodes and extreme AT bias in mitochondrial genes, leading to signal degradation (e.g. Schwarz et al. 2004; Schwarz et al. 2006). To reduce AT bias, the 3rd mtDNA codon position was removed from MP analysis.

BI was implemented with MrBayes v3.1.2 (Huelsenbeck and Ronquist 2001). Codon positions were partitioned for both gene regions, and unlinked for model parameters using default priors, as described by Schwarz et al. (2006). The data set analysed consisted of EF-1α and COI for those *Homalictus* identified as belonging to different putative species. An un-partitioned data set was used in trial analyses but produced minimal support of clades and was subsequently abandoned.

Because of low posterior probability support values, highly related specimens (likely to be conspecifics, see below) were removed from the analysis following trial runs. Markov Chain Monte Carlo (MCMC) chains were run for 3 x 10^6 generations with trees sampled every 500th generation and a burn-in of 1.5 million generations, with post-burnin trees combined across runs to produce a consensus phylogram.

2.3.2 Time of Origin and Speciation

To determine the timing of the *Homalictus* origin, previous studies into halictine radiation which have used representatives from across continents were utilised as a backbone for phylogeny construction (Danforth et al. 2004; Brady et al. 2006). The time of *Homalictus* origin can be determined relative to divergence dates of related taxa and the utilisation of fossil calibration data. For reconstruction of the halictine phylogeny, a dataset of EF-1α for the halictines of Danforth et al. (2004) was compiled and included the *Homalictus* species of this study.

The relationships of major clades in the Halictini tribe established by Danforth et al. (2004) were duplicated as a constraint tree using MacClade v4.06 (Maddison and Maddison 2003). At the node of *Homalictus* within the Australian taxa, this study's phylogeny of *Homalictus* replaced that of Danforth et al. (2004) where a polytomy was allowed. Thus previous work was not reanalysed, but instead utilised to investigate the origin of *Homalictus* without re-exploring relationships among the other taxa. Once constructed, the tree was exported into PAUP* v4.0b (Swofford 1999) where it could be used to constrain the maximum likelihood (ML) analysis.

The data set containing EF-1α for the halictines of Danforth et al. (2004) and *Homalictus* (my data) was analysed in PAUP* v4.0b (Swofford 1999). Before analysis the dataset required the development of an appropriate model of best fit to the data. ModelTest ver 3.06 (Posada and Crandall 1998) was run, applying 56 model types to the data set and a best-fit model of the parameters was chosen for the set

using a log likelihood criterion (Appendix D). The model type was then used to calculate ML branch lengths in the constrained tree using PAUP* v4.0 (Swofford 1999). The MacClade (Maddison and Maddison 2003) constructed tree containing the *Homalictus* species was then loaded as a constraint. The outgroup taxon was set as *Nomioides* (Nomioidinae) which is in the sister sub-family to the Halictinae.

Analysis by ML determined the topological arrangement of the unconstrained *Homalictus* species. ML was used in place of BI as MrBayes v3.1.2 (Huelsenbeck and Ronquist 2001) allows only limited constraint of trees in analyses. ML performed in PAUP* v4.0 (Swofford 1999) will allow all non-*Homalictus* nodes of the tree to be constrained and allow only relationships within *Homalictus* to be determined. MrBayes v3.1.2 (Huelsenbeck and Ronquist 2001) will allow the tree to be constrained to a certain shape; however, the position of the unconstrained clades may differ significantly within the constrained tree. A ML heuristic search was run with topological constraints loaded. Branch swapping was achieved by stepwise addition swapping on best trees with 10 random sequence additions, holding 2 trees at each step for 500 replicates.

The phylogeny of Danforth et al. (2004) was constructed with three gene regions (EF-1α, Wingless, Opsin). As such, the single nuclear (EF-1α) gene region used in my analyses has the potential to conflict with the relationships between taxa based on all three genes. Trial analyses using the single conserved nuclear gene produced zero branch lengths as a result of taxa being highly related and having identical sequences. Negative branch lengths were also produced as the constrained tree shape forced relationships to be expressed for taxa based on the single nuclear gene.

Date estimates for nodes were calculated using the method of penalised likelihood in r8s Version 1.50 (Sanderson 2003). This method uses parametric, nonparametric and semiparametric methods that allow variation in rates along branches using a roughness penalty to relax the assumption of constant rates of evolution to obtain better estimates of rates and times. r8s Version 1.50 (Sanderson 2003) allows topologies of phylograms to be converted to a chronogram showing the relative ages, and in doing this requires some time calibration. This was done by constraining some node times by fixed, minimum or maximum ages using fossil evidence.

The smoothing parameter for the dataset was first calculated within r8s Version 1.50 (Sanderson 2003) and set at 40. Three fossil calibration points ('A', 'B', 'C' in Figure 5) from Danforth et al. (2004) were employed. A minimum divergence between *Augochloropsis metallica* and the *Neocorynura discolor, Augochlorella pomoniella*, and *Megalopta* group was set at 23 myBP (A) because of the existence of fossil Augochlorini species from Baltic amber (Engel 1995; Engel 1996; Engel 2000; Engel and Rightmyer 2000). The divergence of the Caenohalictini bee *Eickwortapis dominicana* from its sister clade, was set at 23 myBP (B) (Michener and Poinar Jr 1996). The last calibration point (C) was based on the Baltic amber fossil of *Electrolictus antiquus* and set to a minimum age of 42 myBP for the Halictini clade (Engel 2001). It should be noted that all three calibrations are minimum ages and that the clades connected to these nodes could be considerably older.

r8s Version 1.50 (Sanderson 2003) was run under Unix in the Terminal application of MacOS X v10.3.9. As r8s Version 1.50 (Sanderson 2003) uses the given root node

maximum age as the actual age, three maximum age analyses were run based on the date and 95% credibility intervals of the Halictinae clade of Danforth et al. (2004). Three runs consisted of; a maximum age of 81 myBP, mean of 70 myBP, and minimum age of 56 myBP.

3. Results

3.1. Location of Origin and Subsequent Dispersals

3.1.1 Maximum Parsimony

The dataset was initially analysed using MP, however due to genetic similarity of many specimens, most nodes lacked good support (i.e. most had bootstrap values of < 85%) (Figure 1). It was found that bifurcations among many specimens could not be resolved, despite removal of the 3rd mtDNA codon position which is known to suffer signal degradation. Some aspects of the topological arrangement differed from that shown in Bayesian analyses (see below). The *Homalictus* root node had low support (43%) with the two subsequent bifurcations having support of less than 75% (73%, 32%). Limited support (< 75%) was evident for most nodes of the *L. (Homalictus) murrayi* and *L. (H.) urbanus* clade. The *L. (H.) urbanus* monophyletic clade had high support (94%), with *L. (H.) murrayi* forming a sister clade of low support (79%). Strangely, *L. (H.) dampieri* C10 was recovered as the sister species to *L. (H.) behri* C7 within the *L. (H.) dotatus* clade, although with low support (59%). *Chilalictus* and *Homalictus* were recovered as sister clades but with limited support (51%). Due to the low support of the majority of clades, little else was drawn from the topology.

3.1.2 Bayesian Inference

Simultaneous Bayesian runs converged after about 3×10^6 generations with standard deviation values for differences in the two simultaneous runs of below 0.1. Posterior

probabilities for tree nodes are given in Figure 2. Comparisons of the MP bootstrap values of Figure 1 indicate higher node support values for the Bayesian analyses. The resultant phylogram is given in Figure 3, with posterior probability values above 95% shown on branch nodes. In comparison to the MP analyses, it is likely that the partitioned Bayesian analyses gave a more credible topology due to fitting a more suitable model to the fast evolving 3rd mtDNA positions as shown in Table 1. It can be seen at the 3rd mtDNA codon there is extreme AT bias with proportions of 47% and 43% respectively. The proportion of invariant sites (4%) for the 3rd mtDNA codon further supports this variability.

To determine the degree of difference indicative of species level divergence, measurement of COI uncorrected 'p' distances (i.e. the total proportion of nucleotide differences between pairs of taxa) were plotted in a histogram (Figure 4). The presence of three peaks is evident. The first appears to coincide with intra-species variability at < 2.5% difference, where most taxa pairs in this peak were keyed to the same species. The highest occurring peak between approximately 6% and 15% appears to coincide with between-species and between-genera differences. The third peak of approximately 18% coincides with divergences of ingroup taxa from the set outgroup. Munasinghe et al. (2003) found an average divergence between species of *Cherax* (Decapoda: Parastacidae) of 7-15 % and 17-23% at the generic level, and found COI to be the most variable of the genes they used. The study of Bucklin et al. (1998) on copepods (Copepoda: Calanoida) also found an average difference of 18% for inter-species divergence in COI.

BI placed the *Homalictus* subgenus as nested within *Chilalictus* with a posterior probability value of 60%. The low support for nodes uniting the sub genera could reflect rapid divergences between these groups or signal degradation if the nodes are very ancient. Also to be considered are the effects of long branch attraction when analyses cover deep divergences, or include clades with very high evolutionary rates. This problem is ameliorated where modelling allows more realistic estimates of base composition and evolutionary rates for 3rd mitochondrial codon positions (Schwarz et al. 2004) but long branch attraction can lead to low or misleading support if not enough ingroup taxa are used (Bergsten 2005). Long branch attraction effects between *Homalictus* and *Chilalictus* may be overcome by including related *Lasioglossum* sub-genera.

Posterior probability values gave high support (> 99%) for all *Homalictus* clades in the BI constructed phylogeny (Figure 3). Due to replicates in morpho-species many genetic groupings indicate specimens with small differences that are likely to be the same species. *L. (H.) dampieri, L. (H.) dotatus, L. (H.) urbanus,* and *L. (H.) murrayi* all contained duplicates of species and arrangements of these duplicates lead to polytomies and/or short branch lengths.

High (> 99%) posterior probability values separated many individual specimens. *L. (H.) nr. dampieri C2,* where the specimen resembled *L. (H.) dampieri* but could not be identified confidently, was separate from the other represented specimens. Such was the case in *L. (H.) dotatus H1, L. (H.) dotatus C8, L. (H.) behri C7, L. (H.) bremerensis H6, L. (H.) blackburni B7, L. (H.) urbanus B1, L. (H.) urbanus B9, and L. (H.) murrayi I10*. Short branch length may indicate further gene regions are

necessary to resolve the small difference between specimens; however, determining relationships within intra-species variation may not be possible.

Paraphyly is evident within the specimens of *L. (H.) dotatus*. The clade of *L. (H.) behri* and *L. (H.) bremerensis* arises within the polytomy with strong support (100%). This provides an interesting arrangement given the collection location of *L. (H.) bremerensis* was in south western Australia and the majority of *L. (H.) dotatus* specimens were collected in northern Australia.

Dispersal patterns can be inferred from geographic locations displayed on tree topology. The southern Australian *L. (H.) murrayi* clade arises within the northern Australia taxa and indicates a likely dispersal from north eastern to south western Australia.

3.2. Time of Origin and Speciation

Using penalised likelihood, the Bayesian phylogram was transformed to produce a chronogram (Figure 5) and dates of divergence for key nodes (1, 2) were estimated. Divergence of *L. (Homalictus)* from *L. (Australictus)* (1) was estimated to have occurred approximately 33.86 myBP (Table 2). This conservative estimation exceeds expected divergence times suggested by Danforth et al. (2004). Dating of the *Homalictus* root node (2) estimates an origin of 25.91 myBP (Table 2). This estimation places the origin approximately within the Late-Oligocene.

Comparison with the work of Danforth et al. (2004) shows concordance in the estimated times of divergence for calibration points of: A, 40.88 myBP; B, 62.42 myBP; C, 58.41 myBP. The halictine clade arises within the Cretaceous/Tertiary boundary, which was also found by Danforth et al. (2004), which is not surprising due to their topology included as a constraint tree for non-*Homalictus* taxa.

Dating analyses required the removal of taxa capable of producing zero branch lengths. *Dinagapostemon sp1*, *Pseudagapostemon pissisi*, and *L. (Lasioglossum) scitulum* were removed from my tree reconstruction of Danforth et al. (2004). In addition to this, duplicate specimens of the phylogram (Figure 3) were pruned. Date estimation is unlikely to have been effected by their removal, as representatives of each major *Homalictus* clade remained. The two additional genes of Danforth et al. (2004) helped resolve genetic relationships across a large geographic range. As such, the topology from my analyses based on the single nuclear EF-1α gene may conflict with Danforth's results, so that the constraint tree could have resulted in negative branch lengths.

A second measure of genetic divergence is indicated in the boxplot of Figure 6 which summarises pairwise ML differences for lineages of the origin of halictine eusociality and the root allodapine node. Known divergence dates for halictine and allodapine lineages (22 myBP, 43 myBP respectively) were used to compare estimated *Homalictus* divergence. The pairwise distances for *Homalictus* taxa whose most recent common ancestor was the *Homalictus* root node were intermediate between the distances found for the allodapine root node and the halictine nodes representing origins of eusociality.

4. Discussion

4.1. Taxonomy

Uncertain *Homalictus* identifications where morphology differed were in some cases, grouped in the phylogeny as genetically similar (e.g. *L. (Homalictus) dampieri*). High posterior probability values (> 99%) along with short branch lengths within these clades suggest the existence of about thirteen *Homalictus* species, of which eight were identified using the key of Walker (1986). For the mitochondrial gene region COI a nucleotide difference of up to approximately 2.5% suggests intra-species variation (Figure 4). In mitochondrial genes a difference of up to 5% has been shown to be indicative of intra-species level divergence (Munasinghe et al. 2003).

The two largest sampled species groups returned interesting assemblages. The '*dampieri*' group contained two distinct phenotypes corresponding to the Queensland and Northern Territory samples respectively, despite Walker (1986) stating variation was minimal in the species. Samples from the coastal north Queensland areas had a darker, almost black base colour, a green mesosoma with gold sheen, and a black gaster suffused with blue and gold. In contrast the Northern Territory specimens had a lighter, brown base colour, a blue (with green tinge) mesosoma with gold sheen, and a brown gaster suffused with blue. Genetically, differences were small, despite consistent morphological differences corresponding to collection location. Absence of a polytomy indicates slight genetic difference, however, is representative of intra specific variation (< 2.5%). Certain specimens such as E6, C10, & E5 show genetic

difference that is potentially higher than intra species variation which may suggest the existence of cryptic species.

Paraphyly of *L. (H.) dotatus* is difficult to interpret. *L. (H.) behri* is restricted to the Northern Territory; however, as *L. (H.) dotatus* is found throughout Australia, genetic distances may represent recent divergence. Moreover, *L. (H.) bremerensis,* a species from south-west Western Australia, is slightly more difficult to explain but may provide support for western dispersal through arid areas. For a widely distributed species such as *L. (H.) dotatus* distribution between the two species may overlap. Further sampling within the Pilbara region of Western Australia may eradicate current lack of support for such movement. Two identified *L. (H.) dotatus* specimens (H1, C8) were separated by high probability values (100%) outside of this paraphyly. Interpretation of this may suggest that while identified species are morphologically distinct, genetically they may remain similar.

Topological arrangement of species relationships often produced uncertainty in identifications. Many specimens within species such as *L. (H.) urbanus* (B1 & B9) showed enough genetic difference to separate them with high support (100%) (Figures 2 and 3). *L. (H.) urbanus* is a cosmopolitan species found throughout Australia, thus intra species genetic variation may be expected. However, well supported nodes (100%) in addition to prominent branch lengths (specimens B1 & B9) may indicate difference above intra-specific variation and thus support for presence of cryptic species.

Identification was complicated due to variation in colouration, but sculpturing characteristics facilitated species distinction. However, where identification was confident, support for genetic relationships was not. This study produced evidence of both, (i) morphologically distinct specimens showing genetic similarity, and (ii) specimens that are similar morphologically but reveal limited genetic difference. This may imply that current keys over estimate the number of extant species when based solely on morphology.

4.2 Dispersal

Plotting the geographic location of the sequenced specimens onto the phylogeny allow some dispersal patterns to be inferred. Although sampling bias may be evident with 70% of *Homalictus* specimens sequenced collected in the Northern Territory, all geographic areas remain represented. The grouping of southern Australia specimens within the northern Australian samples provides support for a north-eastern *Homalictus* origin with subsequent south western dispersal. High diversity of *Homalictus* in north eastern Australia suggests this location of origin is likely. The presence of *L. (H.) bremerensis* within the paraphyletic *L. (H.) dotatus* clade may further suggest western dispersal. *L. (Chilalictus)* specimens were also collected from a number of locations through southern Australia. From the topological arrangement of phylogeny we can also infer there was dispersal to the west in at least one of the *Chilalictus* clades. This may suggest an eastern Australian origin with subsequent western dispersal.

Given a north eastern origin and the proximity of Southeast Asia, it may be assumed that radiation of *Homalictus* into the Papuan/Pacific region occurred at a similar rate as the continental dispersal. However, Southeast Asian species would be necessary for such inference to have any support. Thorough sampling across Sarawak, Malaysia returned no evidence of *Homalictus* in this area. Michener (1979) described **t**he bee fauna of the oriental region as the poorest (89 genera) in the world. Additionally, Liow et al. (2001) reported that the bee fauna in tropical south-east Asia is species poor, despite an extremely high plant species richness (cited in Whitmore 1984). Despite this, *Homalictus* has a wide distribution throughout the Papuan/Pacific region (Michener 1965) and would be expected on Borneo due to their recorded presence on surrounding islands.

4.3. Time of Origin and Speciation

Divergence of the *L. (Homalictus)* lineage from the lineage leading to *L. (Australictus) lithiscum* was estimated at 33 myBP, placing it within the Early-Oligocene, far earlier than what was suggested by Danforth et al. (2004). The *Homalitcus* root node is estimated to have arisen 27 myBP, i.e. in the Late-Oligocene (Gradstein et al. 2004). This has considerable consequences for the expected origin of halictines in Australia due to available passages of dispersal into Australia preceding this date. However, the divergence of extant lineages presented here should not be confused with the origin of *Homalictus*. The origin of *Homalictus* refers to the point at which it became distinct, not when it diverged from other sub-genera or the major homalictus clades diverged.

Root node date estimates are potentially under-estimated due to incomplete sampling. If any unsampled taxa represent older clades the estimated dates in chronogram construction may fail to illustrate previous divergences. This is likely to have occurred in the analysis of Danforth et al. (2004) in which the *Homalictus* root node appears to be recent and divergence from related subgenera appears to be considerably earlier.

As a second measure of genetic divergence, the maximum likelihood distances were calculated (Figure 6). Distances for the node of halictine eusociality origin and the root allodapine node, in which root node age has been previously estimated, were compared with those distances of *Homalictus*. A boxplot of distances is shown in Figure 6, giving medians and inter-quartile ranges. Pairwise distance estimates include the effects of divergences in both lineages descending from the most recent common ancestor to the two species being compared. If the rate of genetic change is linear over time, the time since divergence should be proportional to ML distance divided by two. That means a four-fold difference in distances between two taxa-pairs should correspond, approximately, to a two-fold difference in the divergence ages of the two pairs of taxa being compared.

Topological arrangement of the chronogram of Figure 5 revealed limitation in the single conserved nuclear gene. Conflicting arrangement from Bayesian inference in *L. (H.) urbanus* arising within *L. (H.) murrayi* is likely to be a result of the reduced dataset from the two gene regions that were used in phylogeny construction. EF-1α as a sole measure of genetic divergence is shown to be limited in determining recently diverged taxa.

Dating estimates for deeper nodes are broadly concordant with those of Danforth et al. (2004) and Brady et al (2006). Given that the Halictinae root node age was conservatively set at 70 myBP for a conserved gene region, divergence dates provide strong support for an evolutionary stable communal lineage in the *Lasioglossum*. All *Homalictus* are known to be communal, and as all Australian halictines share this sociality type, communal behaviour is likely to have been present in Australia for an extended period of time.

4.4 Biogeographic Implications

The dates produced here precede the suggested climatic impacts on diversification of halictines proposed by Brady et al. (2006). However, the majority of halictines included in Brady et al.'s (2006) study were from the northern hemisphere or at least not of Australian origin. An estimated origin within the Oligocene pushes the origin of halictines in Australia back earlier than previously found. Danforth et al. (2004) suggested a halictine origin in Australia of approximately 30 myBP. Given the estimated *Homalictus* origin, Australian halictine origin is likely to be much later than this which does not fit the traditional models in biogeography of Gondwanan vicariance or Australian-Laurasian interchange in the Mid-Miocene (Schwarz et al. 2006). Despite the significant water barrier, Southeast Asia remains a possible passage of dispersal due to its proximity. The fragmented landscape and fluctuation of global sea levels throughout the Late-Oligocene (Haq et al. 1987) may have facilitated 'island hopping'. When inferring such dispersal passages, a *Homalictus* origin should not be mistaken for halictine origin in Australia. Especially as southern Australia

potentially inhabits a higher diversity of halictines. Given support for a southern Australia halictine origin may lead to investigation of different passages of dispersal.

Danforth et al. (2004) suggested an African origin for the Halictinae, with subsequent African/South American origin for the Halictini tribe. Bull et al. (2003) suggested the allodapine exoneurine bees dispersed into Australia by either: (i) evolving on the African plate and dispersing into Australia via Asia, or (ii) dispersing over the Indian Ocean, possibly by island-hopping across remnants of the now-submerged Kerguelen Plateau (Krause et al. 1997; Cooper et al. 2001). Given the Early-Oligocene estimates of *Homalictus,* halictine origin must have preceded the Mid-Miocene Laurasian-Australian interchange by a considerable amount of time. Halictine dispersal from Africa to South America was dated at between 70 and 55 myBP, considerably later than the separation of the landmasses. This may support the ability of halictines to cross water barriers, and show that dispersal occurred during this period. It may be suggested that given the opportunity, such as Indian Ocean landmasses, dispersal to the east may have occurred readily.

As an alternative origin, there have been many studies published showing likely dispersal into Australia from Africa from the Late-Eocene to Mid-Oligocene (Bull et al. 2003; Schwarz et al. 2004; Schwarz et al. 2006). Many of these studies produced puzzling results as they suggest distributions that are too recent to be explained by Gondwana vicariance models, where divergences are too ancient for the Mid-Miocene interchange between Laurasia and Australia. The estimated divergence of *Homalictus* may fit this category of uncertainty. Despite the support for a north-eastern Australian origin for *Homalictus* in this studies' phylogeny, there is equal evidence of high

diversity of halictines within southern Australia. Although much of the evidence shows dispersal out of south-east and north-east Australia, these sub-genera may represent early divergence from ancestral halictines. Schwarz et al. (2006) hypothesised an African dispersal of allodapines through island hopping across land masses, such as the Kerguelen Plateau, of the Indian Ocean. This seems plausible for a halictine origin due to the diversity of *Lasioglossum* sub-genera in the southern temperate regions. It has been shown in other taxa (Hopper 1979; Barker et al. 2001; Barker et al. 2007) that following colonisation of Australia, rapid dispersal and speciation has occurred resulting in many endemic groups. Therefore the origin of *Homalictus* in north-eastern Australia should be no indication of a south-east Asian dispersal, and is likely be an outcome of rapid diversification of halictines in Australia.

The halictine genus *Thrincostoma* distribution covers southern and tropical Africa and Madagascar, through to small representatives in southern Asia. Given their presence in southern Africa at the time of origin and early diversification of Halictinae dispersal to southern Asia via the northern hemisphere is possible (Danforth et al. 2004). Given this dispersal into south-east Asia was more recent than the proposed Australian halictine origin, *Thrincostoma* does not demonstrate the ability of bee fauna to cross the Laurasia-Australia water barrier pre-Miocene that would need to be crossed for Australian colonisation in the halictines, but does illustrate a possible route of dispersal from Africa into Australia via southern Asia.

Wherever the ancestral Australian halictines dispersed from, once colonised, radiation of halictines was rapid. Danforth et al. (2004) suggest in relation to the early radiation

of Halictini that as most species are predominantly generalist foragers, the number of available niches to fill would have allowed rapid radiation relative to other more specialist bee faunas. The diversity of Australian habitats and its size as a single land mass would have allowed halictine radiation to occur rapidly.

In summary, *Homalictus* was found to not have a recent divergence, as had been expected. For *Homalictus* the results of this study suggest an origin of north-east Australia with subsequent radiation to both the southwest of Australia supported by the presented phylogeny and to the north-east into the Pacific/Papuan region based on current distributions. Due to the estimated origin of *Homalictus* and consequently, age of halictine origin, dispersal through south-east Asia would have been impeded by a separating water body. Therefore an African dispersal through Indian Ocean land masses is plausible, or at least worth further investigation. Despite this, due to a lack of both African and Indonesian representative taxa, it must be conceded that a south-east Asian origin remains possible despite the Oligocene origin of *Homalictus*.

Acknowledgements

I would like to thank Mike Schwarz for his ongoing assistance with this manuscript. I would like to thank Mike Schwarz, Luke Chenoweth, Sally Harradine, Pelin Kayaalp, Sandra Rehan, and Scott Butler for help with specimen collection. I would also like to thank Mike Gardner and Andrew Craigie for their expertise in genetic methods. I would like to thank Mike Schwarz, Luke Chenoweth, Pelin Kayaalp, Sandra Rehan, Kelly Newton, Sally Harradine, Jaclyn Smith, and Scott Butler for comments and suggestions on this manuscript. Thanks to Trevor Lehmeyer for help with figures.

This work was supported by an Australian Research Council Discovery Grant to M. P. Schwarz (DP0664891).

5. References

Barker, F.K., Barrowclough, G.F. and Groth, J.G. 2001. "A phylogenetic hypothesis for passerine birds: taxonomic and biogeographic implications of an analysis of nuclear DNA sequence data." *Proceedings of the Royal Society of London - Series B Biological Sciences* **269**: 295-308.

Barker, N.P., Weston, P.H., Rutschmann, F. and Sauquet, H. 2007. "Molecular dating of the 'Gondwanan' plant family Proteaceae is only partially congruent with the timing of the break-up of Gondwana." *Journal of Biogeography* **34**: 2012–2027.

Batra, S.W.T. 1966. "Nests and social behaviour of halictine bees of India (Hymenoptera: Halictidae)." *Indian Journal of Entomology* **28**: 375-393.

Benson, D., Karsch-Mizrachi, I., Lipman, D., Ostell, J. and Wheeler, D. 2006. "GenBank." *Nucleic Acids Research* **34**: 16-20.

Bergsten, B. 2005. "A review of long-branch attraction." *Cladistics* **21**: 163-193.

Brady, S.G., Sipes, S., Pearson, A. and Danforth, B.N. 2006. "Recent and simultaneous origins of eusociality in halictid bees." *Proceedings of the Royal Society B-Biological Sciences* **273** 1594: 1643-1649.

Bucklin, A., Bentley, A.M. and Franzen, S.P. 1998. "Distribution and relative abundance of *Pseudocalanus moultoni* and *P. newmani* (Copepoda : Calanoida) on Georges Bank using molecular identification of sibling species." *Marine Biology* **132** 1: 97-106.

Bull, N.J., Schwarz, M.P. and Cooper, S.J.B. 2003. "Phylogenetic Divergence of the Australian allodapine bees (Hymenoptera: Apidae)." *Molecular Phylogenetics and Evolution* **27**: 212-222.

Cooper, A., Lalueza-Fox, C., Anderson, S., Rambaut, A., Austin, J. and Ward, R. 2001. "Complete mitochondrial genome sequences of two extinct moas clarify ratite evolution." *Nature* **409**: 704-707.

Danforth, B.N. 1999. "Phylogeny of the bee genus *Lasioglossum* (Hymenoptera: Halictidae) based on mitochondrial COI sequence data." *Systematic Entomology* **24**: 377-393.

Danforth, B.N. 2002. "Evolution of sociality in a primitively eusocial lineage of bees." *Proceedings of the National Academy of Sciences of the United States of America* **99** 1: 286-290.

Danforth, B.N., Brady, S.G., Sipes, S.D. and Pearson, A. 2004. "Single-copy nuclear genes recover Cretaceous-age divergences in bees." *Systematic Biology* **53** 2: 309-326.

Danforth, B.N., Conway, L. and Ji, S.Q. 2003. "Phylogeny of eusocial *Lasioglossum* reveals multiple losses of eusociality within a primitively eusocial clade of bees (Hymenoptera : Halictidae)." *Systematic Biology* **52** 1: 23-36.

Danforth, B.N. and Ji, S.Q. 2001. "Australian *Lasioglossum* plus *Homalictus* form a monophyletic group: Resolving the "Australian Enigma"." *Systematic Biology* **50** 2: 268-283.

Danforth, B.N., Sauquet, H. and Packer, L. 1999. "Phylogeny of the bee genus *Halictus* (Hymenoptera: Halictidae) based on parsimony and likelihood analysis of nuclear EF-1a sequence data." *Molecular Phylogenetics and Evolution* **13** 3: 605-618.

de Jong, R. 2004. "Phylogeny and biogeography of the genus *Taractrocera* Butler, 1870 (Lepidoptera: Hesperiidae), an example of Southeast Asian-Australian interchange." *Zoologische Mededelingen (Leiden)* **78**: 383-415.

Donovan, B. 1983. "Comparative Biogeography of Native Apoidea of New Zealand and New Caledonia." *GeoJournal* **7** 6: 511-516.

Engel, M.S. 1995. "*Neocorynura electra*, a new fossil bee species from Dominican amber (Hymenoptera: Halictidae)." *Journal of New York Entomological Society* **103**: 317-323.

Engel, M.S. 1996. "New augochlorine bees (Hymenoptera: Halictidae) in Dominican amber, with a brief review of fossil Halictidae." *Journal of Kansas Entomological Society* **69(suppl.)**: 334-345.

Engel, M.S. 2000. "Classification of the bee tribe Augochlorini (Hymenoptera: Halictidae)." *Bulletin of American Museum of Natural History* **250**: 1-89.

Engel, M.S. 2001. "A monograph of the Baltic amber bees and evolution of the Apoidea (Hymenoptera)." *Bulletin of American Museum of Natural History* **259**: 1-192.

Engel, M.S. and Rightmyer, M.G. 2000. "A new augochlorine bee in Tertiary amber from the Dominican Republic (Hymenoptera: Halictidae)." *Apidologie* **31**: 431-436.

Gradstein, F.M., Ogg, J.G. and Smith, A.G. 2004. *A Geologic Time Scale 2004*. Cambridge, UK, Cambridge University Press.

Hall, R. 2002. "Cenozoic geological and plate tectonic evolution of South-East Asia and the South-West Pacific: computer-based reconstructions, model and animations." *Journal of Asian Earth Sciences* **20** 4: 353-431.

Haq, B.U., Hardenbol, J. and Vail, P.R. 1987. "Chronology of fluctuating sea levels since the Triassic." *Science* **235**: 1156-1167.

Hingston, A.B. 1998. "Temporal and spatial variation in abundances of native bee species on an altitudinal gradient in southern Tasmania." *Australian Journal of Zoology* **46**: 497-507.

Hopper, S.D. 1979. "Biogeographical Aspects of Speciation in the Southwest Australian Flora." *Annual Review of Ecology and Systematics* **10** 1: 399-422.

Huelsenbeck, J.P. and Ronquist, F. 2001. "MrBayes: Bayesian Inference of Phylogeny." *Bioinformatics* **17**: 754-755.

Knerer, G. and Schwarz, M.P. 1976. "Halictine Social Evolution: The Australian Enigma." *Science* **194** 445: 445-448.

Krause, D.W., Prasad, G.V.R., Koenigswald, W.v., Sahni, A. and Grine, F.E. 1997. "Cosmopolitanism among Gondwanan Late Cretaceous mammals." *Nature* **390**: 504-507.

Liow, L.H., Sodhi, N.S. and Elmqvist, T. 2001. "Bee Diversity along a Disturbance Gradient in Tropical Lowland Forests of South-East Asia." *Journal of Applied Ecology* **38** 1: 180-192.

Maddison, D.R. and Maddison, W.P. 2003. MacClade version 4.06, Sinauer Associates, Inc.

Michener, C.D. 1958. "The Evolution of Social Behaviour in Bees." *Proceedings of the Tenth International Congree of Entomology* **2**: 441-448.

Michener, C.D. 1960. "Notes on the biology and supposed parthenogenesis of halictine bees from the Australian region." *Journal of the Kansas Entomological Society* **33**: 85–96.

Michener, C.D. 1965. "A Classification of the bees in the Australian and South Pacific regions." *Bulletin of the American Museum of Natural History* **130**: 27-.

Michener, C.D. 1979. "Biogeography of the Bees." *Annals of the Missouri Botanical Garden* **66** 3: 277-347.

Michener, C.D. and Poinar Jr, G. 1996. "The known bee fauna of the Dominican amber." *Journal of Kansas Entomological Society* **69**: 353-361.

Munasinghe, D.H.N., Murphy, N.P. and Austin, C.M. 2003. "Utility of mitochondrial DNA sequences from four gene regions for systematic studies of Australian freshwater crayfish of the genus *Cherax* (Decapoda : Parastacidae)." *Journal of Crustacean Biology* **23** 2: 402-417.

Posada, D. and Crandall, K.A. 1998. "Modeltest: Testing the model of DNA substitution." *Bioinformatics* **14** 9: 817-818.

Rambaut, A. and Charleston, M. 2002. SeAl Sequence Alignment Editor. Oxford, Department of Zoology, University of Oxford.

Sanderson, M.J. 2003. "r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock." *Bioinformatics* **19** 2: 301-302.

Schwarz, M.P., Fuller, S., Tierney, S.M. and Cooper, S.J.B. 2006. "Molecular phylogenetics of the exoneurine allodapine bees reveal an ancient and puzzling dispersal from Africa to Australia." *Systematic Biology* **55** 1: 31-45.

Schwarz, M.P., Richards, M.H. and Danforth, B.N. 2007. "Changing paradigms in insect social evolution: Insights from halictine and allodapine bees." *Annual Review of Entomology* **52**: 127-150.

Schwarz, M.P., Tierney, S.M., Cooper, S.J.B. and Bull, N.J. 2004. "Molecular phylogenetics of the allodapine bee genus *Braunsapis*: A-T bias and heterogeneous substitution parameters." *Molecular Phylogenetics and Evolution* **32**: 110-122.

Smith, B.H. and Wenzel, J.W. 1988. "Pheromonal covariation and kinship in social bee *Lasioglossum zephyrum* (Hymenoptera: Halictidae)." *Journal of Chemical Ecology* **14** 1: 87-94.

Swofford, D.L. 1999. PAUP. Phylogenetic analysis using parsimony (*and other methods). Sunderland, Massachusetts, Sinauer Associates.

Walker, K. 1986. "Revision of the Australian species of the genus *Homalictus* (Cockerell) (Hymenoptera: Halictidae)." *Memoirs of the Museum of Victoria* **47** 2: 105-200.

Walker, K. 1997. "Supplement to a revision of Australian members of the bee genus *Homalictus* (Cockerell) (Hymenoptera: Halictidae)." *Memoirs of the Museum of Victoria* **56** 1: 69-82.

Wcislo, W.T., Wille, A. and Orozco, E. 1993. "Nesting biology of tropical solitary and social sweat bees, *Lasioglossum (Dialictus) figueresi* Wcislo and *L. (D.) aeneiventre* (Friese) (Hymenoptera: Halictidae)." *Insectes Sociaux* **40**: 21-40.

Yan, C. and Kroenke, L.W. 1993. "A plate tectonic reconstruction of the southwest Pacific, 0–100 Ma." *Proceedings of the Ocean Drilling Program Scientific Results* **130**: 697-709.

Table 1: Base frequencies for codon positions of each gene with associated gamma shapes, proportions of invariant sites and relative rates.

Codon Pos.		Nucleotide				Data Shape (alpha)	Proportion of Invariant Sites	Rates
		A	C	G	T			
Nuclear								
1st		0.282	0.196	0.352	0.171	1.030	0.869	1.136
2nd		0.300	0.277	0.159	0.264	18.676	0.883	5.011
3rd		0.203	0.252	0.221	0.324	1.522	0.127	0.196
Mitochondrial								
1st		0.320	0.119	0.182	0.378	0.702	0.441	0.102
2nd		0.238	0.178	0.151	0.434	68.939	0.730	0.026
3rd		0.472	0.087	0.006	0.435	0.875	0.040	0.767

Table 2: Estimated age of nodes for fossil calibration points and *Homalictus* root node and divergence from the most recent common ancestor with upper and lower estimates.

Node	**Age Root Node (myBP)**		
	Upper	**Mean**	**Lower**
	81	70	56
Fossil Calibration Points			
A	39.9	34.5	27.6
B	61.2	52.6	42.1
C	58.2	49.9	42.0
Homalictus			
1	39.5	33.9	28.4
2	29.9	25.9	22.0

Figure 1: Consensus of most parsimonious trees based on a data set with 3rd mitochondrial codon positions excluded. Bootstrap values are based on 500 pseudoreplicates.

Figure 2: Consensus cladogram from a Bayesian analysis with six separately modelled partitions comprising of each EF-1α and COI codon positions. Posterior probabilities and mean branch lengths are derived from 3000 trees taken from 1.5 million generations, sampling every 500th generation. (nr.) = similar appearance but unlikely to be conspecific

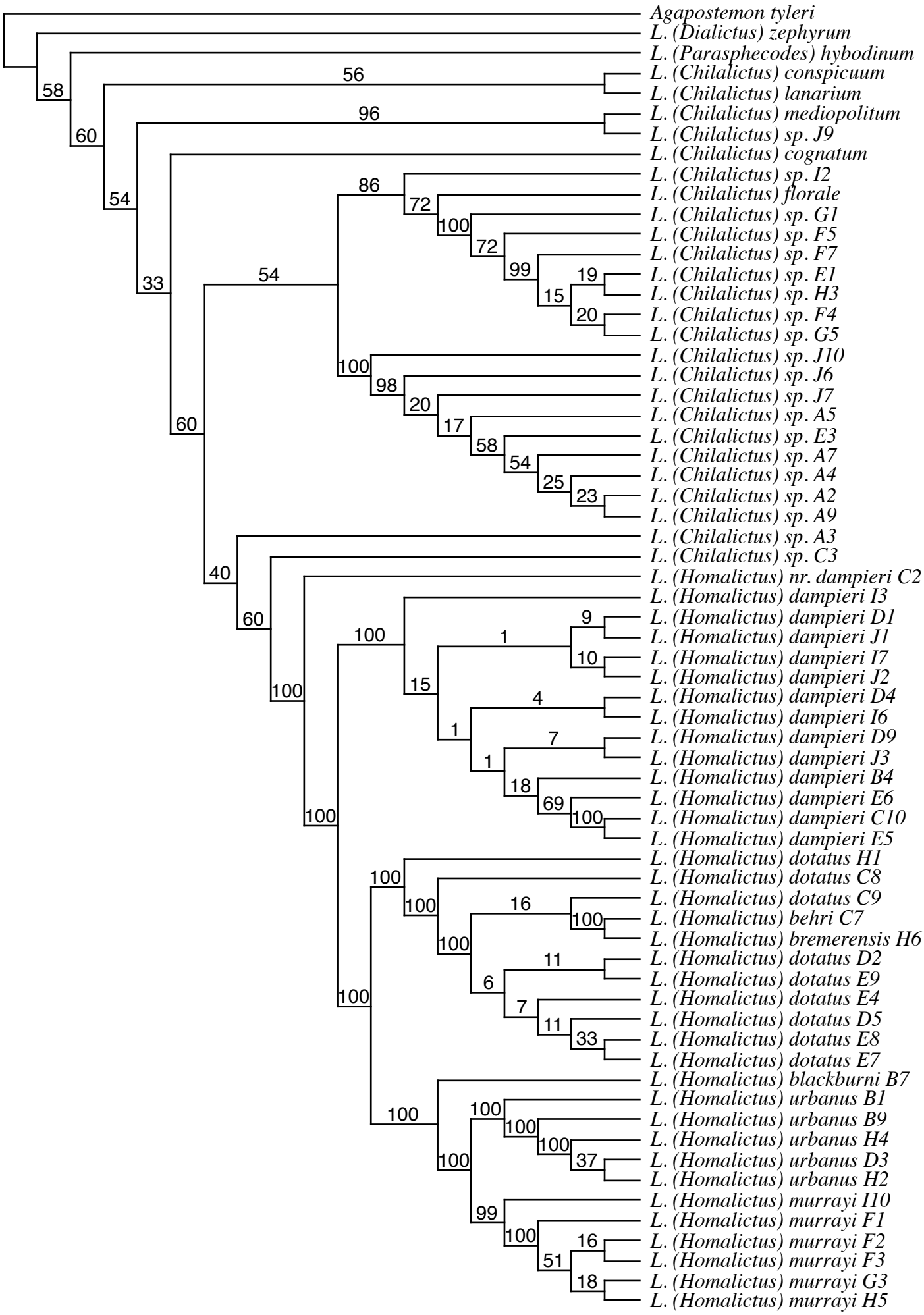
Figure 3: Consensus phylogram from a Bayesian analysis with six separately modelled partitions comprising of each EF-1α and COI codon positions. Nodes of high posterior probability and *Chilalictus* root node are shown. Collection location indicated by colour coded branches.

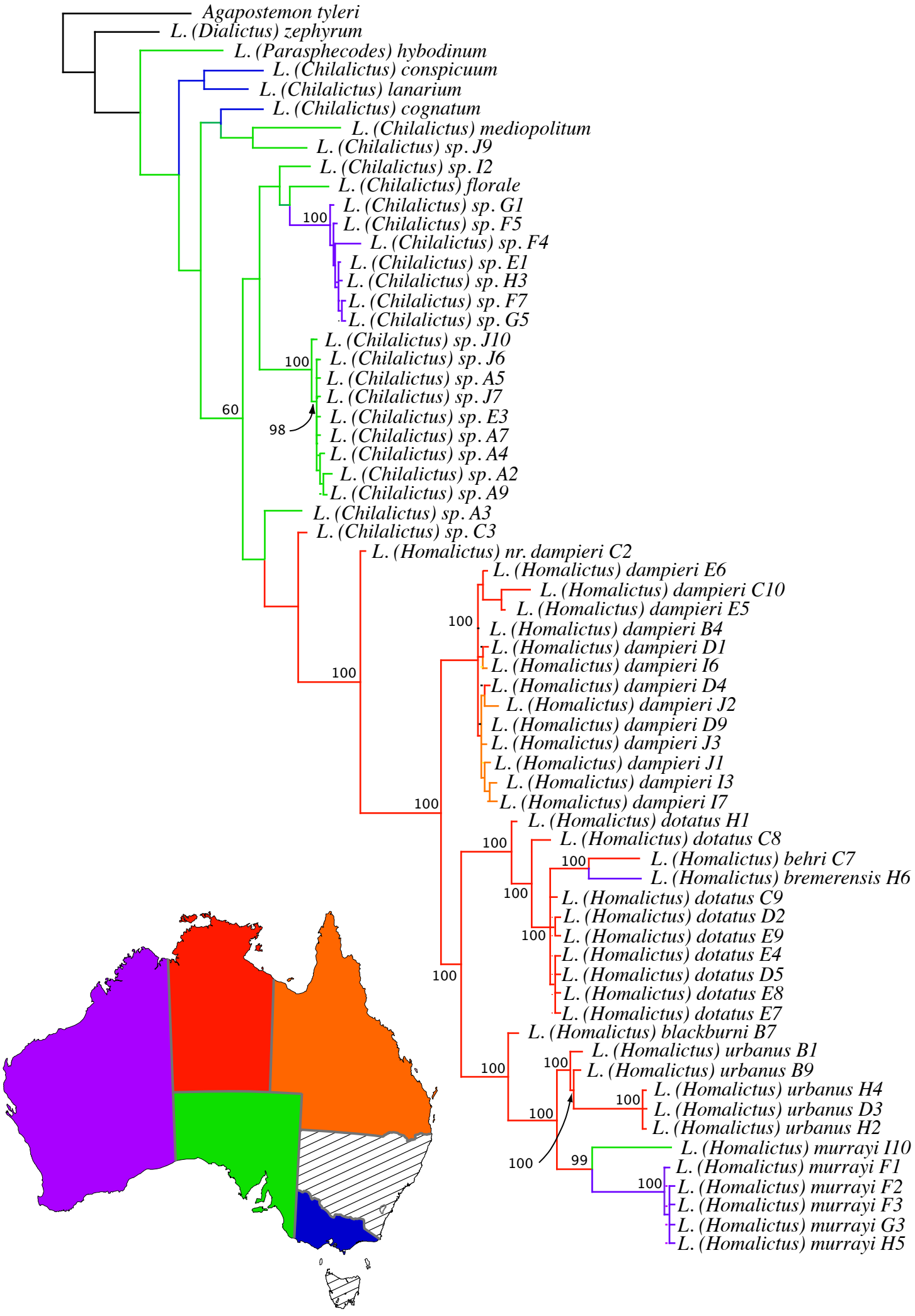
Figure 4: Histogram showing proportion of genetic difference in COI between taxa of phylogram.

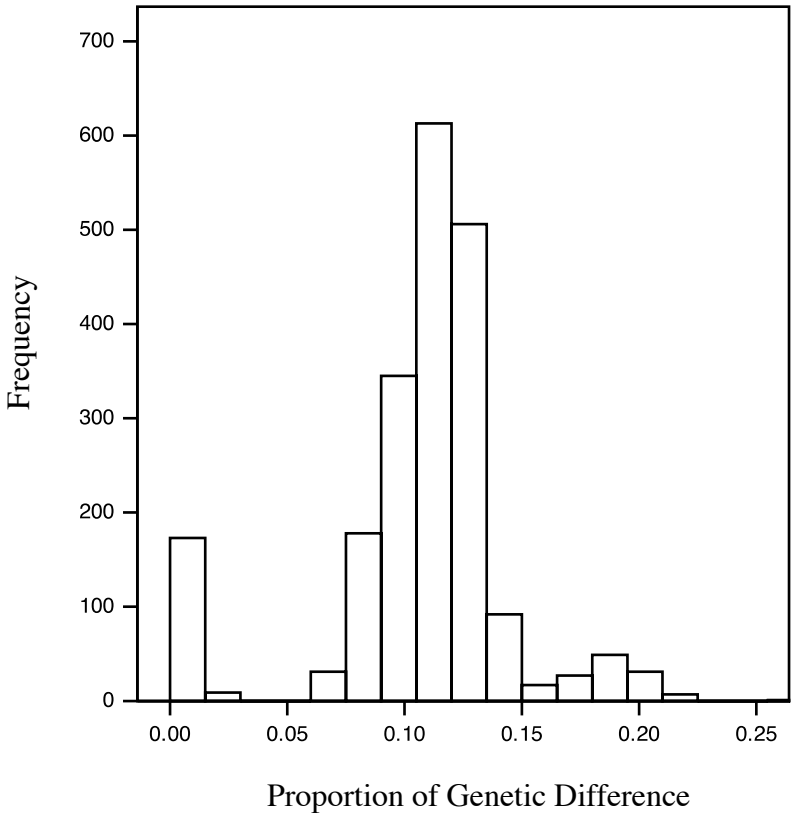
Figure 5: Chronogram constructed through maximum likelihood for *Homalictus*, otherwise constrained. Dating determined through penalised likelihood relaxed clock model. Fossil calibration points indicated by A, B, & C. *Homalictus* divergence from most recent common ancestor indicated by (1). *Homalictus* root node indicated by (2).

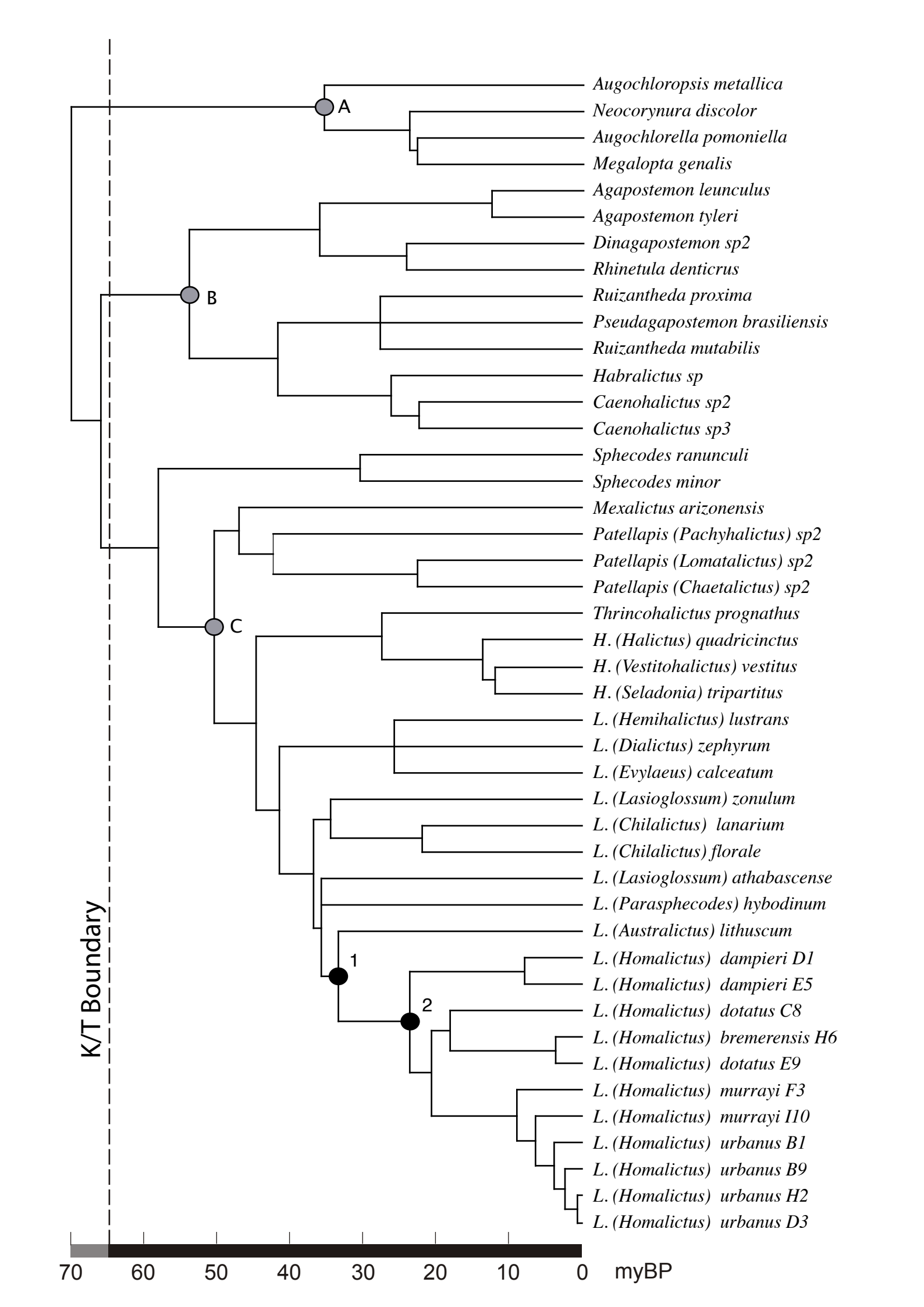
Figure 6: Boxplot giving medians and inter-quartile ranges of compared maximum likelihood distances for the known ages of halictine eusociality origin, root allodapine node, and root *Homalictus* node.

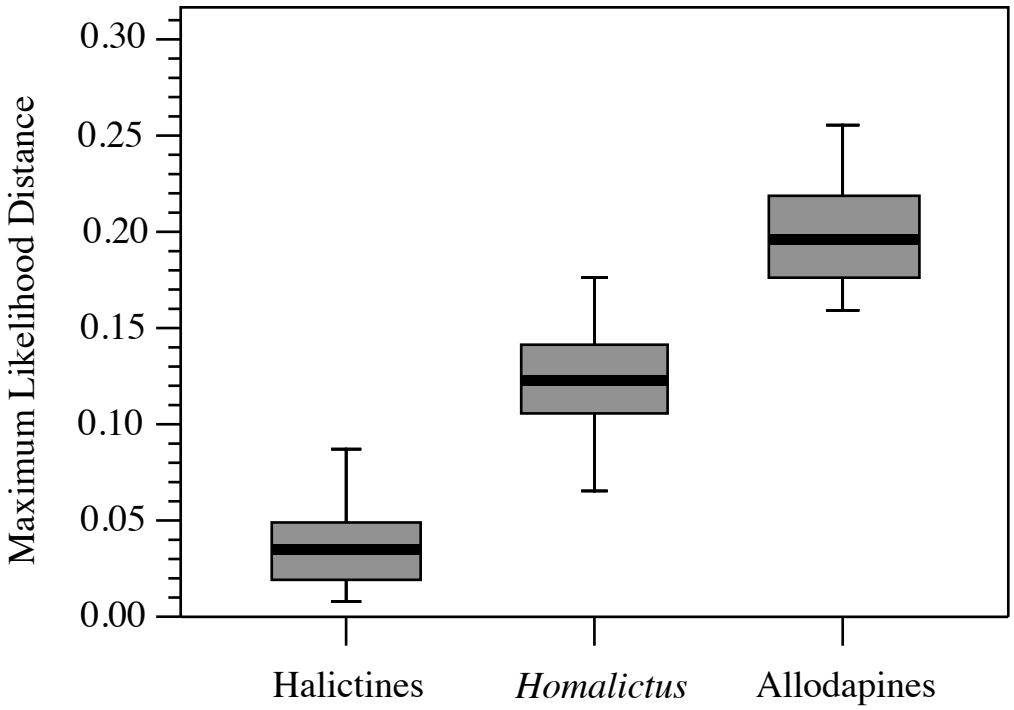

Agapostemon tyleri
L. (Dialictus) zephyrum
L. (Parasphecodes) hybodinum
L. (Chilalictus) cognatum
L. (Chilalictus) sp. J9
L. (Chilalictus) conspicuum
L. (Chilalictus) mediopolitum
L. (Chilalictus) lanarium
L. (Chilalictus) sp. I2
L. (Chilalictus) florale
L. (Chilalictus) sp. F5
L. (Chilalictus) sp. G1
L. (Chilalictus) sp. F4
L. (Chilalictus) sp. E1
L. (Chilalictus) sp. F7
L. (Chilalictus) sp. G5
L. (Chilalictus) sp. H3
L. (Chilalictus) sp. C3
L. (Chilalictus) sp. A3
L. (Chilalictus) sp. J10
L. (Chilalictus) sp. J6
L. (Chilalictus) sp. A5
L. (Chilalictus) sp. J7
L. (Chilalictus) sp. E3
L. (Chilalictus) sp. A2
L. (Chilalictus) sp. A9
L. (Chilalictus) sp. A4
L. (Chilalictus) sp. A7
L. (Homalictus) urbanus B7
L. (Homalictus) dotatus H1
L. (Homalictus) nr. dampieri C2
L. (Homalictus) dampieri I6
L. (Homalictus) dampieri D4
L. (Homalictus) dampieri D9
L. (Homalictus) dampieri I3
L. (Homalictus) dampieri I7
L. (Homalictus) dampieri J3
L. (Homalictus) dampieri D1
L. (Homalictus) dampieri J1
L. (Homalictus) dampieri B4
L. (Homalictus) dampieri J2
L. (Homalictus) dampieri E6
L. (Homalictus) dampieri E5
L. (Homalictus) murrayi F1
L. (Homalictus) murrayi F3
L. (Homalictus) murrayi F2
L. (Homalictus) murrayi G3
L. (Homalictus) murrayi H5
L. (Homalictus) murrayi I10
L. (Homalictus) urbanus B1
L. (Homalictus) urbanus B9
L. (Homalictus) urbanus D3
L. (Homalictus) urbanus H2
L. (Homalictus) urbanusH4
L. (Homalictus) dotatus C8
L. (Homalictus) bremerensis H6
L. (Homalictus) flindersi C7
L. (Homalictus) dampieri C10
L. (Homalictus) dotatus E8
L. (Homalictus) dotatus C9
L. (Homalictus) dotatus D5
L. (Homalictus) dotatus E9
L. (Homalictus) dotatus E7
L. (Homalictus) dotatus D2
L. (Homalictus) dotatus E4
37
23
24
19
47
19
65
59
26
72
51
40
66
35
44
75
37
10
7
16
39
14
22
51
73
65
59
69
0
1
11
6
1
7
1
13
2
54
43
98
22
17
30
89
79
94
83
99
32
35
76
83
59
70
14
7
4
7
17


Agapostemon tyleri
L. (Dialictus) zephyrum
L. (Parasphecodes) hybodinum
L. (Chilalictus) conspicuum
L. (Chilalictus) lanarium
L. (Chilalictus) mediopolitum
L. (Chilalictus) sp. J9
L. (Chilalictus) cognatum
L. (Chilalictus) sp. I2
L. (Chilalictus) florale
L. (Chilalictus) sp. G1
L. (Chilalictus) sp. F5
L. (Chilalictus) sp. F7
L. (Chilalictus) sp. E1
L. (Chilalictus) sp. H3
L. (Chilalictus) sp. F4
L. (Chilalictus) sp. G5
L. (Chilalictus) sp. J10
L. (Chilalictus) sp. J6
L. (Chilalictus) sp. J7
L. (Chilalictus) sp. A5
L. (Chilalictus) sp. E3
L. (Chilalictus) sp. A7
L. (Chilalictus) sp. A4
L. (Chilalictus) sp. A2
L. (Chilalictus) sp. A9
L. (Chilalictus) sp. A3
L. (Chilalictus) sp. C3
L. (Homalictus) nr. dampieri C2
L. (Homalictus) dampieri I3
L. (Homalictus) dampieri D1
L. (Homalictus) dampieri J1
L. (Homalictus) dampieri I7
L. (Homalictus) dampieri J2
L. (Homalictus) dampieri D4
L. (Homalictus) dampieri I6
L. (Homalictus) dampieri D9
L. (Homalictus) dampieri J3
L. (Homalictus) dampieri B4
L. (Homalictus) dampieri E6
L. (Homalictus) dampieri C10
L. (Homalictus) dampieri E5
L. (Homalictus) dotatus H1
L. (Homalictus) dotatus C8
L. (Homalictus) dotatus C9
L. (Homalictus) behri C7
L. (Homalictus) bremerensis H6
L. (Homalictus) dotatus D2
L. (Homalictus) dotatus E9
L. (Homalictus) dotatus E4
L. (Homalictus) dotatus D5
L. (Homalictus) dotatus E8
L. (Homalictus) dotatus E7
L. (Homalictus) blackburni B7
L. (Homalictus) urbanus B1
L. (Homalictus) urbanus B9
L. (Homalictus) urbanus H4
L. (Homalictus) urbanus D3
L. (Homalictus) urbanus H2
L. (Homalictus) murrayi I10
L. (Homalictus) murrayi F1
L. (Homalictus) murrayi F2
L. (Homalictus) murrayi F3
L. (Homalictus) murrayi G3
L. (Homalictus) murrayi H5


Agapostemon tyleri
L. (Dialictus) zephyrum
L. (Parasphecodes) hybodinum
L. (Chilalictus) conspicuum
L. (Chilalictus) lanarium
L. (Chilalictus) cognatum
L. (Chilalictus) mediopolitum
L. (Chilalictus) sp. J9
L. (Chilalictus) sp. I2
L. (Chilalictus) florale
L. (Chilalictus) sp. G1
L. (Chilalictus) sp. F5
L. (Chilalictus) sp. F4
L. (Chilalictus) sp. E1
L. (Chilalictus) sp. H3
L. (Chilalictus) sp. F7
L. (Chilalictus) sp. G5
L. (Chilalictus) sp. J10
L. (Chilalictus) sp. J6
L. (Chilalictus) sp. A5
L. (Chilalictus) sp. J7
L. (Chilalictus) sp. E3
L. (Chilalictus) sp. A7
L. (Chilalictus) sp. A4
L. (Chilalictus) sp. A2
L. (Chilalictus) sp. A9
L. (Chilalictus) sp. A3
L. (Chilalictus) sp. C3
L. (Homalictus) nr. dampieri C2
L. (Homalictus) dampieri E6
L. (Homalictus) dampieri C10
L. (Homalictus) dampieri E5
L. (Homalictus) dampieri B4
L. (Homalictus) dampieri D1
L. (Homalictus) dampieri I6
L. (Homalictus) dampieri D4
L. (Homalictus) dampieri J2
L. (Homalictus) dampieri D9
L. (Homalictus) dampieri J3
L. (Homalictus) dampieri J1
L. (Homalictus) dampieri I3
L. (Homalictus) dampieri I7
L. (Homalictus) dotatus H1
L. (Homalictus) dotatus C8
L. (Homalictus) behri C7
L. (Homalictus) bremerensis H6
L. (Homalictus) dotatus C9
L. (Homalictus) dotatus D2
L. (Homalictus) dotatus E9
L. (Homalictus) dotatus E4
L. (Homalictus) dotatus D5
L. (Homalictus) dotatus E8
L. (Homalictus) dotatus E7
L. (Homalictus) blackburni B7
L. (Homalictus) urbanus B1
L. (Homalictus) urbanus B9
L. (Homalictus) urbanus H4
L. (Homalictus) urbanus D3
L. (Homalictus) urbanus H2
L. (Homalictus) murrayi I10
L. (Homalictus) murrayi F1
L. (Homalictus) murrayi F2
L. (Homalictus) murrayi F3
L. (Homalictus) murrayi G3
L. (Homalictus) murrayi H5
100
100
100
60
98
100
100
100
100
100
100
100
100
100
100
100
100
99
100


Frequency
700
600
500
400
300
200
100
0
0.00
0.05
0.10
0.15
0.20
0.25
Proportion of Genetic Difference


A
B
C
1
2
K/T Boundary
Augochloropsis metallica
Neocorynura discolor
Augochlorella pomoniella
Megalopta genalis
Agapostemon leunculus
Agapostemon tyleri
Dinagapostemon sp2
Rhinetula denticrus
Ruizantheda proxima
Pseudagapostemon brasiliensis
Ruizantheda mutabilis
Habralictus sp
Caenohalictus sp2
Caenohalictus sp3
Sphecodes ranunculi
Sphecodes minor
Mexalictus arizonensis
Patellapis (Pachyhalictus) sp2
Patellapis (Lomatalictus) sp2
Patellapis (Chaetalictus) sp2
Thrincohalictus prognathus
H. (Halictus) quadricinctus
H. (Vestitohalictus) vestitus
H. (Seladonia) tripartitus
L. (Hemihalictus) lustrans
L. (Dialictus) zephyrum
L. (Evylaeus) calceatum
L. (Lasioglossum) zonulum
L. (Chilalictus) lanarium
L. (Chilalictus) florale
L. (Lasioglossum) athabascense
L. (Parasphecodes) hybodinum
L. (Australictus) lithuscum
L. (Homalictus) dampieri D1
L. (Homalictus) dampieri E5
L. (Homalictus) dotatus C8
L. (Homalictus) bremerensis H6
L. (Homalictus) dotatus E9
L. (Homalictus) murrayi F3
L. (Homalictus) murrayi I10
L. (Homalictus) urbanus B1
L. (Homalictus) urbanus B9
L. (Homalictus) urbanus H2
L. (Homalictus) urbanus D3
70
60
50
40
30
20
10
0
myBP


Maximum Likelihood Distance
0.30
0.25
0.20
0.15
0.10
0.05
0.00
Halictines
Homalictus
Allodapines